An Automated Soil Line Identification Routine for Remotely Sensed Images

Garey A. Fox,* G. J. Sabbagh, S. W. Searcy, and C. Yang

ABSTRACT

The soil line is a linear relationship between the near-infrared (NIR) and red (R) reflectance of bare soil as characterized by slope and intercept parameters. Vegetation indices use soil line parameters extensively in crop growth analyses. Research indicates that the soil line can be related to site-specific soil conditions within a field, especially organic C content. This relationship may provide a means for directing soil sampling. However, these soil and crop growth remotely sensed predictions require accurate estimates of soil line parameters. Determining soil line parameters by manually extracting reflectance characteristics of bare soil pixels can be cumbersome. This research proposes an automated soil line identification routine capable of deriving soil line parameters from bare soil or vegetated remotely sensed images. The automated routine estimates soil line parameters by deriving a set of minimum NIR digital numbers across the R band range. Pixels that contradict soil line theory are removed through an iterative process. The routine was evaluated using bare soil images of two fields in the Midwest USA and 15 multispectral digital video images of South Texas grain sorghum fields dominated by vegetated cover. This research compared soil line parameters derived from the automated routine to actual soil line parameters obtained by extracting R and NIR digital numbers from identifiable bare soil pixels within the images and also by manually inspecting plots of R versus NIR digital numbers for all pixels within an image. The routine performed reasonably well in matching the estimated actual soil line parameters with minimal adjustment between images.

The soil line is a linear relationship between the NIR and R reflectance of bare soil originally discovered by Richardson and Wiegand (1977):

$$NIR = \beta_1 R + \beta_0 \quad [1]$$

where β_1 is the soil line slope and β_0 is the intercept. The soil line extends from an upper region of bright soils with high reflectance in both the R and NIR bands (Point B in Fig. 1) to a lower region consisting of darker soils with low R and NIR reflectance or digital number (Point A in Fig. 1). The soil line for a particular soil type "…results from the combined variations of its surface status characterized by its roughness and moisture" (Baret et al., 1993). Jasinski and Eagleson (1989) demonstrate that three unique soil lines result by varying soil type, moisture content, and roughness. A global soil line representing all soil types is not possible because such a line will only be linear in portions of the entire range due to variations caused by different soil conditions (soil type, moisture content, organic matter content, soil roughness, etc.).

Recent research indicates that the R and NIR reflectance or image intensity can be used as a mechanism for deriving soil properties from remotely sensed images and guiding soil sampling (Fox and Sabbagh, 2002; Wilcox et al., 1994). Percentage of organic mater is known to relate to the reflectance of bare soils (Frazier, 1989; Henderson et al., 1989; Baumgardner et al., 1985). Estimates of organic matter content are essential for determining soil fertility and soil physical properties (Tisdale et al., 1995). Organic matter content influences soil structure and correspondingly aeration and infiltration (Pepper, 1996). Fox and Sabbagh (2002) indicate that soil organic matter is correlated to a pixel's location along the soil line (i.e., a function of the R and NIR reflectance or image intensity) and may be a larger influence on soil line parameters than soil type or roughness for Mollisols. Fox and Sabbagh (2002) utilize the soil line concept to develop predictive relationships between the amount of organic matter within the surface horizon of the soil profile and intensity in the R and NIR bands. The soil line Euclidean distance (SLED) technique relates a pixel's Euclidean distance of the R and NIR to the R and NIR reflectance for the bottommost point on the soil line. This Euclidean distance measurement is shown to correlate well to surface organic matter content in two midwestern cornfields. However, deriving accurate soil line parameters is critical to developing these predictive relationships. Furthermore, research is still needed to determine how the combined influences of moisture content, organic C, and other soil environmental conditions relate to observed reflectance. An automated routine for determining soil line parameters can provide a consistent methodology for investigating these relationships.

Soil line parameters are also used extensively in deriving estimates of vegetation growth through vegetation indices. Remotely sensed measurements of vegetation are made through the use of biomass–vegetation index relationships. Vegetation indices are based on the principle that significant differences exist in the reflections of healthy vegetation, senesced vegetation, and dry bare soil as a function of wavelength (Jensen, 1996). These indices make it possible to integrate information on the growth of vegetation throughout the growing season. Vegetation indices are commonly related to biomass through leaf area index (LAI). Leaf area index measures the leaf area of the vegetation per unit area of soil surface and is important in considering photosynthetic

G.A. Fox, 208 Carrier Hall, Dep. of Civil Engineering, Univ. of Mississippi, University, MS 38677-1848; G.J. Sabbagh, Bayer CropScience, 17745 S. Metcalf, Stilwell, KS 66085; S.W. Searcy, Dep. of Biological & Agricultural Engineering, Texas A&M University, College Station, TX 77843-2117; C. Yang, USDA-ARS, Kika de la Garza, Subtropical Agricultural Research Unit, 2413E Hwy 83, Weslaco, TX 78596. Received 27 Mar. 2003. *Corresponding author (gafox@olemiss.edu).

Published in Soil Sci. Soc. Am. J. 68:1326–1331 (2004).

677 S. Segoe Rd., Madison, WI 53711 USA

Abbreviations: LAI, leaf area index; NDVI, normalized difference vegetation index; NIR, near-infrared; PVI, perpendicular vegetation index; R, red; SAVI, soil adjusted vegetative index; TSAVI, transformed soil adjusted vegetative index.

activity and yield (Campbell, 1996). A widely used relation between vegetation indices and LAI is Beer's Law:

$$\frac{\text{IPAR}}{\text{PAR}} = 1 - \exp(-k\text{LAI}) \quad [2]$$

IPAR is intercepted solar radiation, PAR is photosynthetically active radiation, and k is the light extinction coefficient (Flenet et al., 1996).

Two common classes of indices are the subject of considerable research: (i) the normalized difference vegetation index (NDVI); and (ii) the soil line related indices, such as the perpendicular vegetation index, PVI (Richardson and Wiegand, 1977); soil adjusted vegetation index, SAVI (Huete, 1988); and transformed soil adjusted vegetation index, TSAVI (Baret et al., 1989). NDVI, declared as an intrinsic index by Rondeaux et al. (1996), is the most commonly used vegetation index:

$$\text{NDVI} = \frac{(\text{NIR} - \text{R})}{(\text{NIR} + \text{R})} \quad [3]$$

NIR is the reflectance or digital number in the near-infrared band and R is the reflectance or digital number in the red band. The sensitivity of the NDVI to soil background and atmospheric effects is the reason for interest in new indices, which utilize the soil line concept (Rondeaux et al., 1996). These indices are related to the soil line concept developed by Richardson and Wiegand (1977). Richardson and Wiegand originally propose the PVI, a Euclidean distance measurement of the pixel's NIR and R reflectance from the soil line (Point C and D in Fig. 1). As shown below, PVI can be expressed as a linear combination of the NIR and R reflectance and as a function of the slope (β_1) and intercept (β_0) of the soil line (Baret and Guyot, 1991):

$$\text{PVI} = \frac{1}{\sqrt{\beta_1^2 + 1}}(\text{NIR} - \beta_1\text{R} - \beta_0) \quad [4]$$

Experiments with the PVI demonstrate that points corresponding to the same canopy do not migrate along lines parallel to the soil line when the soil brightness is changing (Baret and Guyot, 1991). Therefore, darker soils tend to result in lower values for the vegetation index than brighter soils with the same amount of vegetative cover (i.e., the slope of the percentage vegetation lines are not parallel with the soil line). Huete (1988) proposes the SAVI derived from NDVI. The SAVI can be expressed in terms of the NIR and R reflectance and also a constant (L) to minimize soil brightness influences, according to the following:

$$\text{SAVI} = \frac{(\text{NIR} - \text{R})}{(\text{NIR} + \text{R} + L)}(1 + L) \quad [5]$$

However, Baret and Guyot (1991) express inconsistencies in this vegetation index as well, stating that SAVI is an exact solution only for soil lines in which the slope is exactly unity and the intercept is exactly zero. Baret et al. (1989) propose the TSAVI to alleviate this inconsistency:

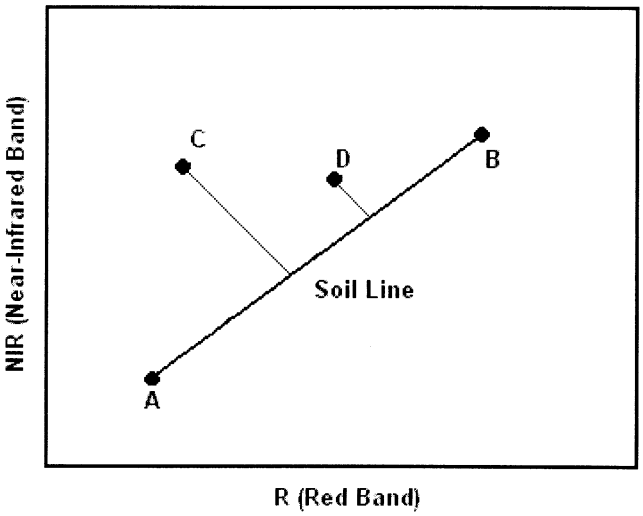

Fig. 1. Soil line concept demonstrating the observed linear relationship between red (R) and near-infrared (NIR) reflectance or image intensity of bare soil (from Fox et al., 2003).

$$\text{TSAVI} = \frac{\beta_1(\text{NIR} - \beta_1\text{R} - \beta_0)}{\beta_1\text{NIR} + \text{R} - \beta_1\beta_0 + X(1 + \beta_1^2)} \quad [6]$$

X is a constant, usually assumed 0.08 for true reflectance values; TSAVI equals zero for bare soil and is close to 0.70 for highly vegetated surfaces.

More recent research focuses on the most appropriate index for crop investigations. Normalized difference vegetation index is the most widely used and accepted because of its simplistic form. Rondeaux et al. (1996) suggest reasons why the soil line related indices have encountered less frequent adoption in remote sensing. First, soil line related indices possess a more complex formulation (e.g., necessity of deriving soil line parameters). Determining soil line parameters is difficult due to the inability to develop a global soil line (Baret et al., 1993). Soil-line parameters must be determined for each research site. Also, a lack of convincing evidence exists for improved assessments of vegetation by the soil line related indices compared to NDVI. Rondeaux et al. (1996) conclude that NDVI is sensitive to the soil background reflectance and that the soil line related indices appear to be more reliable and less noisy than the NDVI. Baret and Guyot (1991) conclude that soil line related indices also are less noisy, especially for low LAIs (LAI $<$ 2–3). However, Wiegand et al. (1991) report no significant differences between the two different categories of vegetation indices.

Soil parameter analysis and vegetation crop growth investigations with remotely sensed images require accurate estimates of soil line parameters. For example, percentage of organic matter in the upper horizon of the soil profile is shown to directly correlate to Euclidean distance along the soil line (Fox and Sabbagh, 2002). To quantify a pixel's Euclidean distance from the bottom-most point on a soil line, accurate estimates of the soil line slope and intercept are required. Furthermore, vegetation indices are sensitive to soil line parameters (Baret et al., 1993). Sensitivity of a vegetation index with respect to the soil line slope or intercept can be

quantified by taking the derivative of the index with respect to the soil line parameters (i.e., $\partial(\text{PVI})/\partial\beta_1$ and $\partial(\text{PVI})/\partial\beta_0$). In general, the maximum sensitivity of PVI to β_0 is considerably less than the sensitivity of PVI to β_1. For instance, the sensitivity of PVI to β_0 is only dependent on the slope of the soil line. Perpendicular vegetative index is most sensitive to β_0 when β_1 is small (i.e., soil lines with small slopes) such that soil conditions are uniform across the field (i.e., equivalent soil type, low gradient in moisture content and a uniform organic matter content). The sensitivity of PVI to β_1 is dependent on NIR, R, β_1, and β_0. Perpendicular vegetative index sensitivity to β_1 increases as NIR and R increase. The sensitivity of PVI to the soil line slope increases as β_0 (soil line intercept) decreases. Similar results in terms of the sensitivity with respect to the soil line slope versus the soil line intercept are expected for the other soil line related vegetation indices.

As demonstrated above, more detailed research is needed in studying soil property and crop growth relationships with remotely sensed images and soil line parameters. Deriving soil line parameters by extracting reflectance from manually determined bare soil pixels within an image could be cumbersome. The objective of this research is to develop an automated system capable of deriving parameters for the soil line of any image. Such an automated system will provide an efficient and consistent methodology for deriving soil line parameters for both soils and crop investigations. A description of the routine is presented. Then, the routine is evaluated by comparing soil line parameters derived by the automated routine to estimated actual soil line parameters from both bare soil and vegetated dominated images.

AUTOMATED SOIL LINE DEVELOPMENT ROUTINE

The automated routine estimates soil line parameters by deriving a set of minimum NIR digital numbers across the R band range based on a user-defined bandwidth. The bandwidth determines the number of minimum NIR values selected for each R-value present in the data. For example, a bandwidth of unity selects the single-most minimum NIR value for each R-value. These R and corresponding minimum NIR values form digital number or reflectance pairs: (R_i $NIR^*_{i,n}$), where NIR* is the minimum near-infrared value, i = 0...255 and n = bandwidth.

The routine calculates an initial soil line using a subset of these (R_i, $NIR^*_{i,n}$) pairs. The routine determines the subsets by calculating quartiles based on the range along the R band. Quartiles correspond to 0, 25, 50, 75, and 100% of this range. The routine then calculates soil lines and regression coefficients from the (R_i, $NIR^*_{i,n}$) pairs in the quartile ranges. The automated routine uses the soil line with the largest coefficient of determination as the initial soil line with parameters β_1 and β_0. Even though this initial soil line may include pixels not dominated by bare soil, this estimate guides the automated routine in subsequent steps.

To remove (R_i, $NIR^*_{i,n}$) pairs not dominated by bare soil, the routine makes use of an iterative procedure to check for *outlier* pixels. Outlier pixels are potentially vegetated pixels that are commonly selected on the lower portion of the soil line, unique environmental attributes in the image such as standing water or shadows of clouds, or pixels with vegetation uniquely stressed compared with other pixels due to the lack of soil water or nutrients. For example, soils with higher organic matter content, soil fertility, and correspondingly higher moisture holding capacity dominate pixels in the lower end of the soil line. Except under extremely saturated conditions, vegetation growth from these pixels will predominate at the beginning of the growing season due to the favorable crop growth conditions within the soil. Therefore, the number of bare soil pixels in this range becomes limited. Vegetation dominates the minimum NIR values selected by the routine. The routine needs to remove these vegetative-dominated pixels to estimate the soil line.

The iterative procedure begins with the leftmost (R_i, $NIR^*_{i,n}$) pair and selects m pairs. Initial values of m are usually on the order of five to ten. The routine derives soil line parameters ($\beta_{1,m}$, $\beta_{0,m}$) for this iterative subset using linear regression techniques. Next, the routine compares the initial soil line to the soil line from this iterative subset. If the soil lines are different, then the routine calculates the vertical deviation, ε, of each (R_i, $NIR^*_{i,n}$) pair in the subset to the subset's soil line:

$$\varepsilon_{i,n} = NIR^*_{i,n} - (\beta_0 + \beta_1 R_i) \quad [7]$$

The routine removes the (R_i, $NIR^*_{i,n}$) pair with the maximum vertical deviation, max ($\varepsilon_{i,n}$).

The routine then moves to the next m pairs along the R band range and follows the same procedure for removing pixels. When the entire R band range is processed, the routine calculates a new initial soil line based on the remaining (R_i, $NIR^*_{i,n}$) pairs. The subset size, m, increases by one unit and the routine follows the same iterative procedure again. The subset size increases until either the soil line parameters converge or until the entire R range is encompassed within one subset. The routine performs a linear regression on the remaining (R_i, $NIR^*_{i,n}$) pairs to derive the final soil line parameters.

MATERIALS AND METHODS

This research used images from five different fields to evaluate the routine by comparing estimated actual soil lines and automated soil lines. Bare soil images of two midwestern cornfields (*Zea mays* L.) (Field 1 and 2) were acquired using a digital camera system with area array technology (digital charged coupled devices) and a spatial resolution of 0.5 m. Field 1 was a 0.32-km^2 (80-acre) field located in Buchanan County, IA and Field 2 was a 0.43-km^2 (106-acre) field located in Fremont County, IA. Fox and Sabbagh (2002) utilized the images of Fields 1 and 2 for investigating the relationship between soil line parameters and percentage of organic matter in the soil surface. The aerial images provided four bands of data: blue (400–500 nm), green (G)(500–600 nm), R (600–700 nm), and NIR (700–1000 nm). The digital imaging system acquired the bare soil images of Field 1 and 2 on 10 May 1998 and 1 June 1998, respectively.

This research also utilized vegetated images of three South Texas grain sorghum [*Sorghum bicolor* (L.) Moench] fields (Fields A, S, and N). A multispectral digital video imaging system captured the images. Yang and Anderson (1999) used the images for Fields S and N to study spatial plant growth variability. The video system provided three bands of data: NIR, 845 to 857 nm; R, 625 to 635 nm; and G, 555 to 565 nm. The five images of Field A were during different crop development stages within the 1998-growing season. Three images of Fields S and N were during the 1995-growing season, and two images of Fields S and N were during the 1996-growing season.

Table 1. Image acquisition dates and reported crop development stages for Fields A, S, and N.

	Image date	Crop stage
	Field A	
1998	April 22nd	Pre-boot to boot
	May 11th	Boot to half-bloom
	May 18th	Bloom to soft-dough
	May 29th	Soft-dough to hard-dough
	June 16th	Physiological maturity
	Field S	
1995	May 5th	Boot to half-bloom
	May 16th	Half-bloom to soft-dough
	June 14th	Hard-dough
1996	May 10th	Half-bloom to soft-dough
	May 20th	Soft-dough to hard-dough
	Field N	
1995	May 5th	Boot
	May 16th	Half-bloom to soft-dough
	June 14th	Hard-dough
1996	May 10th	Half-bloom to soft-dough
	May 20th	Soft-dough to hard-dough

Table 1 shows the specific image acquisition dates and reported prevalent crop growth stages. The videos imaging system captured all grain sorghum images at least 10 d after irrigation. Therefore, no significant changes in soil intensity due to soil moisture variations were expected.

The automated routine derived soil lines using digital numbers with the realization that soil lines in reflectance space will be different. Users must realize such differences in soil line parameters in the application of biomass–vegetation index relationships to quantify crop growth and development. Digital numbers record the intensity of electromagnetic energy measured for each pixel. To convert these numbers into reflectance values, low and high reflectance homogeneous Lambertian targets of know spectral characteristics must be included in the image (Sabins, 1987).

This research used image-processing software (ArcView) to derive soil lines by visually identifying bare-soil pixels within the image. The image processing software allowed digital numbers of bare-soil pixels to be extracted directly from the image. These bare soil pixels provided an "estimated actual" soil. This research compared the estimated actual soil lines to "estimated" soil lines based on visual displays of NIR versus R-values for all pixels within the grid files following a Gram-Schmidt procedure (Crist, 1983). The estimated soil line indicated whether selected bare-soil pixels characterized the soil line along its entire R-band range. Additional bare soil pixels selected from the image were combined with prior bare-soil pixels to derive an "estimated actual" soil line for each image. This research compared estimated actual and automated soil lines based on confidence intervals around the estimated actual soil line. Automated and actual soil lines were not significantly different if the 95% confidence intervals around the estimated actual soil line, derived from bare-soil pixels extracted directly from the image, bound the automated soil line.

RESULTS AND DISCUSSION

The automated routine first estimated soil lines for the bare soil images of the two midwestern cornfields. The routine easily identified bare soil pixels and the soil line parameters. However, bare soil images may have multiple bare soil pixels for a given R-value due to differences in soil characteristics such as moisture content, roughness, and percentage of organic matter. To address these issues, an extended bandwidth, n, was used to select a greater number of minimum NIR pixels. The routine provided improved results when using a bandwidth of five to ten units, and an initial subset size, $m = 5$. Bandwidth was directly influenced by the spatial resolution of the image: as spatial resolution increased, the routine required a larger bandwidth to capture the soil line parameters. Results using $n = 5$ and $m = 5$ for the 0.5-m resolution images are shown in Table 2.

Table 2. Comparison of automated and estimated actual soil lines. All results are with a bandwidth, $n = 1$, and initial interval size, $m = 5$, unless otherwise indicated.

			Slope, β1	Intercept, β0
Field 1	5/10/98	Estimated actual	0.76	42.65
		Automated†	0.75	41.63
Field 2	6/1/98	Estimated actual	0.57	40.61
		Automated†	0.55	43.12
Field A	4/22/98	Estimated actual	0.65	−22.72
		Automated	0.70	−25.63
	5/11/98	Estimated actual	0.68	−18.42
		Automated	0.70	−13.04
	5/18/98	Estimated actual	0.79	−36.04
		Automated	0.73	−20.87
	5/29/98	Estimated actual	0.19	61.57
		Automated	0.16	64.21
	6/16/98	Estimated actual	0.56	25.44
		Automated	0.64	21.09
Field S	5/5/95	Estimated actual	0.67	−65.75
		Automated‡	0.70	−71.24
	5/16/95	Estimated actual	0.83	−50.83
		Automated	0.78	−42.99
	6/14/95	Estimated actual	0.59	−0.31
		Automated	0.54	1.25
	5/10/96	Estimated actual	0.62	8.63
		Automated	0.68	−1.13
	5/20/96	Estimated actual	0.48	−8.07
		Automated	0.55	−17.75
Field N	5/5/95	Estimated actual	0.57	−20.21
		Automated§	0.63	−16.87
	5/16/95	Estimated actual	0.81	−30.88
		Automated	0.81	−28.62
	6/14/95	Estimated actual	0.73	−9.37
		Automated§	0.76	−13.35
	5/10/96	Estimated actual	0.51	27.61
		Automated‡	0.55	27.73
	5/20/96	Estimated actual	0.47	−5.02
		Automated	0.42	1.61

† Using a bandwidth, $n = 5$, and initial subset size, $m = 5$.
‡ Using a bandwidth, $n = 1$, and initial subset size, $m = 10$.
§ Using a bandwidth, $n = 1$, and initial subset size, $m = 15$.

Table 2 also shows results for images of Fields A, S and N. Bare soil dominated none of the images. As such, it was much more difficult to determine soil line parameters. The images did possess bare soil pixels around the edges of the fields, which contributed significantly to the success of the routine. Table 2 compares the estimated actual soil line parameters with the automated soil line parameters derived using a bandwidth, $n = 1$, and initial subset size, $m = 5$. Figures 2 through 5 illustrate results from the automated soil line development routine, including the NIR versus R plot for all pixels within the grid files, (R, NIR) pairs selected by the automated soil line identification routine, and comparison of the automated soil line and estimated actual soil line. For example, the estimated actual and automated soil line parameter values were not significantly different with $n = 1$ and $m = 5$, as the automated soil line fell within the boundaries of the 95% confidence intervals around the estimated actual soil line for the 22 Apr. 1998 image of Field A. The routine provided similar

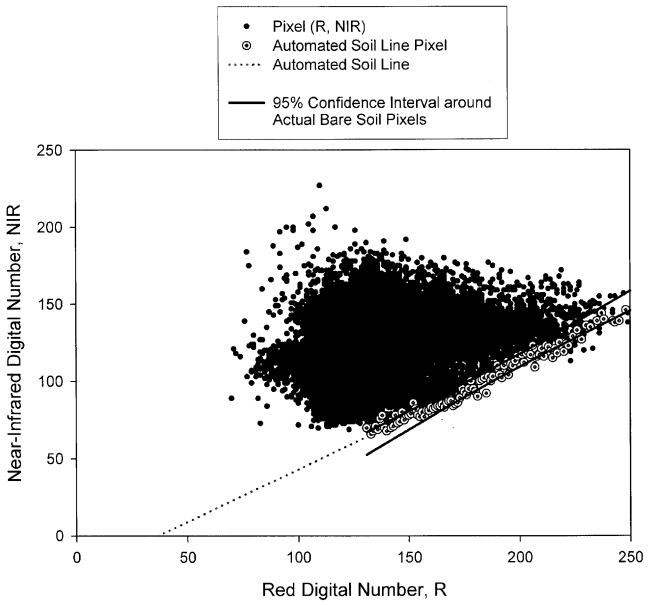

Fig. 2. Comparison of estimated actual and automated soil lines for 22 Apr. 1998 image of Field A.

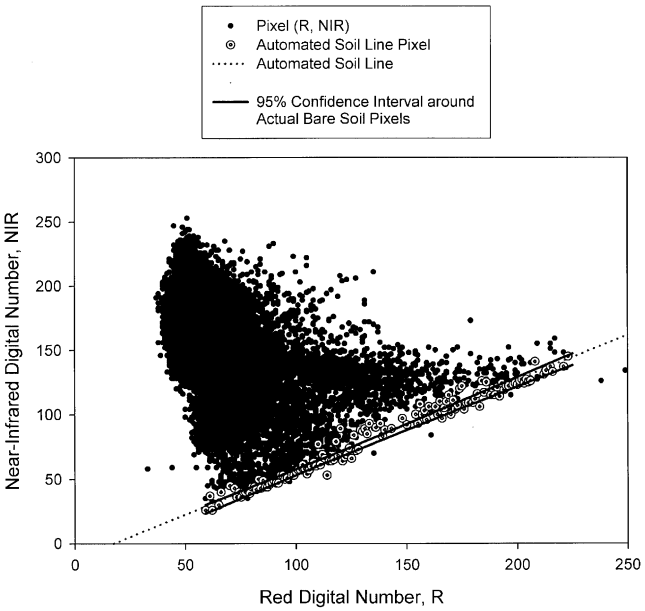

Fig. 4. Comparison of estimated actual and automated soil lines for 16 May 1995 image of Field S.

results for the May 11th image of Field A, 16 May 1995 image of Field S, and the 16 May 1995 image of Field N as illustrated in Fig. 3, 4, and 5, respectively.

However, in several images, the routine required increases in the initial subset size to obtain automated soil lines within the 95% confidence intervals of the estimated actual soil lines. Automated soil lines fell within the 95% confidence intervals when the initial subset size was increased to $m = 15$ in the 5 May 1995 and 14 June 1995 images of Field N. For most images, an initial subset size, $m = 5$, produced reasonable estimates of the actual soil line parameters. However, four images required an expanded initial subset size to obtain a reasonable match. Based on a comparison of these images, the need for the extended subset did not correspond to image conditions, soil characteristics or vegetation growth patterns. Further investigation is required with additional images to develop concrete predictable relationships. Increasing the bandwidth size beyond $n = 1$ deteriorated results in vegetated dominated images as more pixels dominated by vegetated cover needed to be removed.

Additionally, a sixth image of Field A, acquired on 15 Apr. 1998, was not included in the analysis. This image possessed constricted digital number ranges along both the R and NIR bands, appearing as a single cluster with little variation. When attempting to apply the automated routine to such an image, the routine reported

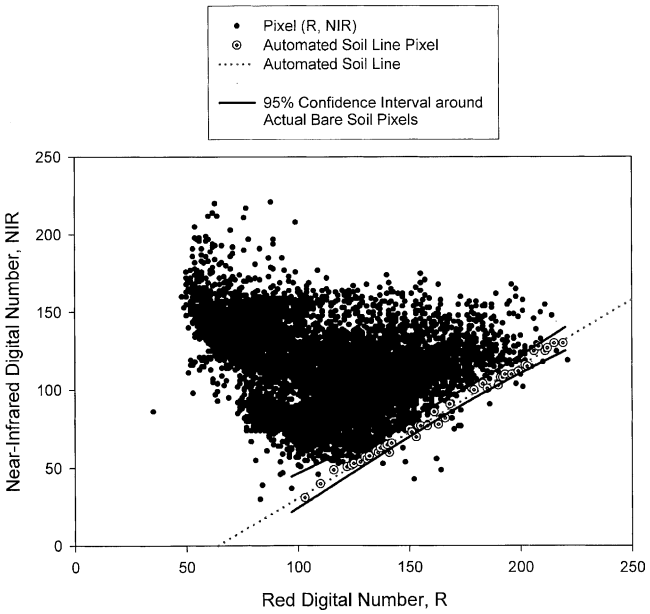

Fig. 3. Comparison of estimate actual and automated soil lines for 11 May 1998 image of Field A.

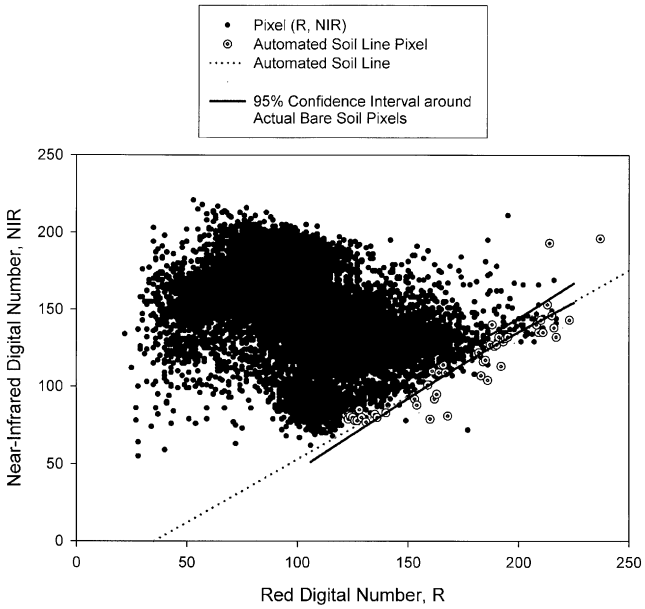

Fig. 5. Comparison of estimated actual and automated soil lines for 16 May 1995 image of Field N.

unreasonable soil line parameters (i.e., no soil line slope). Mechanisms for detecting such image conditions and highlighting the need for user intervention need to be developed. Users must recognize unreasonable soil line parameters based on visual investigations of the R versus NIR digital numbers for pixels within an image. To overcome the requirement of at least a quality assurance visual inspection of such a plot, additional mechanisms need to be included into the automated routine, such as restrictions in terms of the width of the range along both bands or possibly through a moment-major axis relationship.

CONCLUSIONS

This research proposed and evaluated an automated soil line identification routine. The routine begins by deriving a set of minimum NIR digital numbers across the R band range based on a user-defined bandwidth. The user-defined bandwidth determines the number of minimum NIR values selected for each red band value present in the data. The routine derives an initial soil line from these R and minimum NIR data pairs and then uses an iterative procedure to remove pixels not corresponding to soil line theory. This research used digital aerial images during bare soil conditions of two Midwestern cornfields and sets of multispectral digital video images of South Texas grain sorghum fields to evaluate the routine. Analysis involved comparing soil lines derived by the automated soil line identification routine to estimated actual soil lines based on digital numbers in the R and NIR bands of identified bare soil pixels in the image and plots of the red versus near-infrared digital numbers for all pixels within the image.

The routine adequately determined the soil line parameters for all images based on a comparison of the automated soil line with 95% confidence intervals around the estimated actual soil line. The success of the automated routine was dependent on the user-defined bandwidth parameter and an initial subset size for the iteration process. In general, for bare soil images, the routine required a typical bandwidth of five to ten units to select multiple NIR values for each R band value present in the data. For images dominated by vegetation, a bandwidth of unity produced the most reasonable results. The initial subset size did not impact results for the bare soil images. However, the routine required increased initial subset size in 4 of the 15 vegetation-dominated images. The need for an extended initial subset size did not correspond to specific soil or environmental conditions. Therefore, further research is needed to determine when expansion is required. In general, results for the automated soil line identification routine are promising. The ability to automatically derive soil line parameters without manual investigation of bare soil pixels within the images allows soil parameter relationships and soil line related vegetation indices to be more easily investigated. Future research should focus on understanding the capabilities and improving the methodology of the automated soil line routine.

ACKNOWLEDGMENTS

The authors acknowledge the financial support of the EPA STAR Fellowship Program, Grant Number U91-5329, for supporting this research. The authors also acknowledge three anonymous reviewers for suggestions that significantly improved the manuscript.

REFERENCES

Baret, F., and G. Guyot. 1991. Potentials and limitations of vegetation indices for LAI and APAR assessment. Remote Sens. Environ. 35: 161–173.

Baret, F., G. Guyot, and D. Major. 1989. TSAVI: A vegetation index which minimizes soil brightness effects on LAI and APAR estimation. p. 1355–1359. *In* 12th Canadian symposium on remote sensing and IGARSS'90. Vol. 4. 10–14 July 1989. Vancouver, Canada. Geoscience and Remote Sensing Society of Institute of Electrical and Electronics Engineers, IEEE, Piscataway, NJ.

Baret, F., S. Jacquemond, and J.F. Hanocq. 1993. The soil line concept in remote sensing. Remove Sens. Rev. 7:65–82.

Baumgardner, M.F., L.F. Silva, L.L. Biehe, and E.R. Stoner. 1985. Reflectance properties of soils. Adv. Agron. 38:1–44.

Campbell, J.B. 1996. Introduction to remote sensing. 2nd ed. Guilford Press, New York.

Crist, E.P. 1983. The thematic mapper tasseled cap: A preliminary formulation. p. 357–363. *In* Proceedings of machine processing of remotely sensed data. Laboratory for Applications in Remote Sensing, West Lafayette, IN.

Flenet, F., J.R. Kiniry, J.E. Board, M.E. Westgate, and D.C. Reicosky. 1996. Row spacing effects on light extinction coefficients of corn, sorghum, soybean, and sunflower. Agron. J. 88:185–190.

Fox, G.A., and G.J. Sabbagh. 2002. Estimation of soil organic matter from red and near-infrared remotely sensed data using a soil line Euclidean distance technique. Soil Sci. Soc. Am. J. 66:1922–1928.

Fox, G.A., G.J. Sabbagh, and S.W. Searcy. 2003. Radiometric normalization of multi-temporal images using the soil line transformation technique. Trans. ASAE 46:851–859.

Frazier, B.E. 1989. Use of Landsat Thematic Mapper band ratios for soil investigations. Adv. Space Res. 9:155–158.

Henderson, T.L., A. Szilagyi, M.F. Baumgardner, C.T. Chen, and D.A. Landgrebe. 1989. Spectral band selection for classification of soil organic matter content. Soil Sci. Soc. Am. J. 53:1778–1784.

Huete, A.R. 1988. A soil adjusted vegetation index (SAVI). Remote Sens. Environ. 25:295–309.

Jasinski, M.F., and P.S. Eagleson. 1989. The structure of red-infrared scattergrams of semivegetated landscapes. IEEE Trans. Geosci. Remote Sens. 27:441–451.

Jensen, J.R. 1996. 2nd ed. Introductory image processing: A remote sensing perspective. Prentice Hall, Upper Saddle River, NJ.

Pepper, I.L. 1996. Abiotic characteristics of soil. p. 9–18. *In* I.L. Pepper et al. (ed.) Pollution science. Academic Press, NY.

Richardson, A.J., and C.L. Wiegand. 1977. Distinguishing vegetation from soil background information. Photogramm. Eng. Remote Sens. 43:1541–1552.

Rondeaux, G., M. Steven, and F. Baret. 1996. Optimization of soil-adjusted vegetation indices. Remote Sens. Environ. 55:95–107.

Sabins, F.F., Jr. 1987. Remote sensing: Principles and interpretation. W.H. Freeman, New York.

Tisdale, S.L., W.L. Nelson, and J.D. Beaton. 1995. Soil fertility and fertilizers. 4th ed. MacMillian Publ. Co., New York.

Wiegand, C.L., A.J. Richardson, D.E. Escobar, and A.H. Gerbermann. 1991. Vegetation indices in crop assessments. Remote Sens. Environ. 35:105–119.

Wilcox, C.H., B.E. Frazier, and S.T. Ball. 1994. Relationship between soil organic carbon and Landsat TM data in Eastern Washington. Photogramm. Eng. Remote Sens. 60:777–781.

Yang, C., and G.L. Anderson. 1999. Airborne videography to identify spatial plant growth variability for grain sorghum. Precis. Agric. 1:67–79.